Stat 602 Homework 2

Kiegan Rice and Nate Garton

Due 3/4/2019

Problem 1

Problem 4.3

A table of RMSEs for 8-fold cross-validation can be found in Table 1.

Table 1: 8-fold CV Model RMSEs for Problem 4.3

Model	**8-fold CV RMSE**
None	36643.38
Size	27863.72
Fireplace	29008.41
Bsmt Bath	34502.36
Land	33652.75
Size, Fireplace	22886.91
Size, Bsmt Bath	26534.75
Size, Land	25380.03
Fireplace, Bsmt Bath	27230.51
Fireplace, Land	27002.92
Bsmt Bath, Land	32853.91
Size, Fireplace, Bsmt Bath	23034.4
Size, Fireplace, Land	21994.35
Fireplace, Bsmt Bath, Land	27659.4
Size, Fireplace, Bsmt Bath, Land	22347.5

The lowest RMSE appears to be for the full model which includes Size, Fireplace, Bsmt Bath, and Land.

Problem 4.4

Part a

```
## k-Nearest Neighbors
##
## 146 samples
##   9 predictor
##   2 classes: '1', '2'
##
## Pre-processing: centered (9), scaled (9)
## Resampling: Cross-Validated (10 fold)
## Summary of sample sizes: 132, 131, 132, 131, 132, 132, ...
## Resampling results across tuning parameters:
##
##   k   Accuracy   Kappa
##    1  0.8000000  0.6001655
##    2  0.7861905  0.5723620
##    3  0.8280952  0.6569934
##    4  0.7804762  0.5627531
##    5  0.7947619  0.5902949
```

```
##    6  0.7680952  0.5378942
##    7  0.7600000  0.5223285
##    8  0.7538095  0.5104451
##    9  0.7880952  0.5794132
##   10  0.7676190  0.5399746
##   11  0.7733333  0.5521184
##   12  0.7733333  0.5498348
##   13  0.7938095  0.5915763
##   14  0.7800000  0.5630535
##   15  0.7866667  0.5758542
##   16  0.7723810  0.5480266
##   17  0.7652381  0.5335155
##   18  0.7652381  0.5337483
##   19  0.7728571  0.5504010
##   20  0.7514286  0.5064151
##
## Accuracy was used to select the optimal model using the largest value.
## The final value used for the model was k = 3.
```

Based on 10-fold cross-validation, the accuracy is estimated to be highest when $k = 4$.

Part b

The classification rule $\hat{y} = 2I\left[t(x) \geq k/2\right] + I\left[t(x) < k/2\right]$ is equivalent to $\hat{y} = 2I\left[t(x)/k \geq 1/2\right] + I\left[t(x)/k < 1/2\right]$. Here, $t(x)/k$ is an estimate of $E\left[\text{type} = 2|x\right] = P(\text{type} = 2|x)$. Thus, this is directly an estimate of a posterior class probability which accounts for the prior class probabilities. This means that no modification of kNN needs to be made to account for differing prior probabilities.

Problem 5.1

The following are decompositions using the matrix

$$X = \begin{bmatrix} 2 & 4 & 7 & 2 \\ 4 & 3 & 5 & 5 \\ 3 & 4 & 6 & 1 \\ 5 & 2 & 4 & 2 \\ 1 & 3 & 4 & 4 \end{bmatrix}$$

Part a

The QR decomposition of X is:

$$Q = \begin{bmatrix} -0.27 & 0.57 & 0.77 & 0.04 \\ -0.54 & -0.07 & -0.13 & 0.56 \\ -0.4 & 0.37 & -0.35 & -0.71 \\ -0.67 & -0.5 & 0.1 & -0.13 \\ -0.13 & 0.53 & -0.5 & 0.4 \end{bmatrix}, \quad R = \begin{bmatrix} -7.42 & -6.07 & -10.25 & -5.53 \\ 0 & 4.15 & 5.99 & 2.28 \\ 0 & 0 & 1.06 & -1.23 \\ 0 & 0 & 0 & 3.57 \end{bmatrix}$$

Thus, the basis for $C(X)$ will be the columns of the matrix Q.

The Singular Value Decomposition of X is:

$$U = \begin{bmatrix} -0.5 & 0.53 & 0.16 & -0.66 \\ -0.5 & -0.59 & 0.13 & -0.11 \\ -0.46 & 0.49 & -0.25 & 0.65 \\ -0.39 & -0.33 & -0.68 & -0.09 \\ -0.37 & -0.17 & 0.65 & 0.35 \end{bmatrix}, \quad D = \begin{bmatrix} 16.58 & 0 & 0 & 0 \\ 0 & 3.78 & 0 & 0 \\ 0 & 0 & 3.38 & 0 \\ 0 & 0 & 0 & 0.55 \end{bmatrix},$$

$$V = \begin{bmatrix} -0.4 & -0.43 & -0.8 & 0.12 \\ -0.44 & 0.3 & 0.18 & 0.83 \\ -0.71 & 0.45 & 0.04 & -0.54 \\ -0.38 & -0.72 & 0.57 & -0.06 \end{bmatrix}$$

Thus, the basis for $C(X)$ will be the columns of the matrix U.

Part b

We can find the eigen (spectral) decomposition of $X'X$ by calculating the eigenvalues, D^2, and the eigenvectors, which will be the columns of V:

$$D^2 = \begin{bmatrix} 274.94 \\ 14.32 \\ 11.44 \\ 0.3 \end{bmatrix}, \quad V = \begin{bmatrix} -0.4 & -0.43 & -0.8 & 0.12 \\ -0.44 & 0.3 & 0.18 & 0.83 \\ -0.71 & 0.45 & 0.04 & -0.54 \\ -0.38 & -0.72 & 0.57 & -0.06 \end{bmatrix}$$

We can find the eigen (spectral) decomposition of XX' by calculating the eigenvalues, D^2, and the eigenvectors, which will be the columns of U (rather than V as for $X'X$):

$$D^2 = \begin{bmatrix} 274.94 \\ 14.32 \\ 11.44 \\ 0.3 \end{bmatrix}, \quad V = \begin{bmatrix} -0.5 & 0.53 & 0.16 & -0.66 \\ -0.5 & -0.59 & 0.13 & -0.11 \\ -0.46 & 0.49 & -0.25 & 0.65 \\ -0.39 & -0.33 & -0.68 & -0.09 \\ -0.37 & -0.17 & 0.65 & 0.35 \end{bmatrix}$$

Part c

The best $rank = 1$ approximation to X will be $X^{*1} = U_1 diag(d_1) V_1'$, where U_1 is a matrix with the first column of U, and V_1 is a matrix with the first column of V:

$$X^{*1} = U_1 diag(d_1) V_1' = \begin{bmatrix} -0.5 \\ -0.5 \\ -0.46 \\ -0.39 \\ -0.37 \end{bmatrix} \times \begin{bmatrix} 16.58 \end{bmatrix} \times \begin{bmatrix} -0.4 & -0.44 & -0.71 & -0.38 \end{bmatrix}$$

$$= \begin{bmatrix} 3.35 & 3.61 & 5.89 & 3.11 \\ 3.38 & 3.64 & 5.94 & 3.14 \\ 3.08 & 3.31 & 5.4 & 2.85 \\ 2.63 & 2.83 & 4.62 & 2.43 \\ 2.45 & 2.64 & 4.31 & 2.27 \end{bmatrix}$$

The best $rank = 2$ approximation to X will be $X^{*2} = U_2 diag(d_1, d_2) V_2'$, where U_2 is a matrix with the first two columns of U, and V_2 is a matrix with the first two columns of V:

$$X^{*2} = U_2 diag(d_1, d_2) V_2' = \begin{bmatrix} -0.5 & 0.53 \\ -0.5 & -0.59 \\ -0.46 & 0.49 \\ -0.39 & -0.33 \\ -0.37 & -0.17 \end{bmatrix} \times \begin{bmatrix} 16.58 & 0 \\ 0 & 3.78 \end{bmatrix} \times \begin{bmatrix} -0.4 & -0.44 & -0.71 & -0.38 \\ -0.43 & 0.3 & 0.45 & -0.72 \end{bmatrix}$$

$$= \begin{bmatrix} 2.49 & 4.2 & 6.78 & 1.66 \\ 4.35 & 2.97 & 4.95 & 4.75 \\ 2.28 & 3.86 & 6.23 & 1.51 \\ 3.16 & 2.46 & 4.07 & 3.33 \\ 2.73 & 2.44 & 4.02 & 2.75 \end{bmatrix}$$

Part d

Note that

$$\tilde{X} = \begin{bmatrix} -1 & 0.8 & 1.8 & -0.8 \\ 1 & -0.2 & -0.2 & 2.2 \\ 0 & 0.8 & 0.8 & -1.8 \\ 2 & -1.2 & -1.2 & -0.8 \\ -2 & -0.2 & -1.2 & 1.2 \end{bmatrix}$$

The Singular Value Decomposition of $\tilde{X}$ is:

$$\tilde{U} = \begin{bmatrix} -0.57 & 0.17 & 0.29 & -0.6 \\ 0.51 & 0.11 & 0.7 & 0.21 \\ -0.5 & -0.25 & -0.09 & 0.69 \\ 0.34 & -0.69 & -0.32 & -0.33 \\ 0.22 & 0.65 & -0.57 & 0.03 \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} 3.83 & 0 & 0 & 0 \\ 0 & 3.38 & 0 & 0 \\ 0 & 0 & 2.01 & 0 \\ 0 & 0 & 0 & 0.48 \end{bmatrix}$$

$$\tilde{V} = \begin{bmatrix} 0.34 & -0.81 & 0.45 & 0.18 \\ -0.37 & 0.18 & 0.26 & 0.88 \\ -0.58 & 0.03 & 0.68 & -0.45 \\ 0.64 & 0.56 & 0.52 & 0 \end{bmatrix}$$

Therefore, the principal component directions will be the columns of $\tilde{V}$, the principal components will be the inner products $z_j = \langle x_i, v_j \rangle$, the columns of:

$$z = \begin{bmatrix} -2.19 & 0.56 & 0.57 & -0.29 \\ 1.95 & 0.39 & 1.4 & 0.1 \\ -1.91 & -0.84 & -0.18 & 0.33 \\ 1.3 & -2.32 & -0.65 & -0.16 \\ 0.85 & 2.21 & -1.14 & 0.02 \end{bmatrix}$$

The "loadings" of the first principal component are the first column of the matrix $\tilde{V}$,

$$\tilde{V}_1 = \begin{bmatrix} 0.34 \\ -0.37 \\ -0.58 \\ 0.64 \end{bmatrix}$$

Part e

The best $rank = 1$ approximation to $\tilde{X}$ will be $\tilde{X}^{*1} = \tilde{U}_1 diag(\tilde{d}_1), \tilde{V}_1$, where $\tilde{U}_1$ is a matrix with first column of $\tilde{U}$ and $\tilde{V}_1$ is a matrix with the first column of $\tilde{V}$:

$$\tilde{X}^{*1} = \tilde{U}_1 diag(\tilde{d}_1)\tilde{V}_1' = \begin{bmatrix} -0.57 \\ 0.51 \\ -0.5 \\ 0.34 \\ 0.22 \end{bmatrix} \times \begin{bmatrix} 3.83 \end{bmatrix} \times \begin{bmatrix} 0.34 & -0.37 & -0.58 & 0.64 \end{bmatrix}$$

$$= \begin{bmatrix} -0.75 & 0.81 & 1.26 & -1.41 \\ 0.67 & -0.72 & -1.12 & 1.26 \\ -0.66 & 0.71 & 1.1 & -1.23 \\ 0.45 & -0.48 & -0.75 & 0.84 \\ 0.29 & -0.31 & -0.49 & 0.55 \end{bmatrix}$$

The best $rank = 2$ approximation to $\tilde{X}$ will be $\tilde{X}^{*2} = \tilde{U}_2 diag(\tilde{d}_1, \tilde{d}_2)\tilde{V}_2'$, where $\tilde{U}_2$ is a matrix with the first two columns of $\tilde{U}$, and $\tilde{V}_2$ is a matrix with the first two columns of $\tilde{V}$:

$$\tilde{X}^{*2} = \tilde{U}_2 diag(\tilde{d}_1, \tilde{d}_2)\tilde{V}_2' = \begin{bmatrix} -0.57 & 0.17 \\ 0.51 & 0.11 \\ -0.5 & -0.25 \\ 0.34 & -0.69 \\ 0.22 & 0.65 \end{bmatrix} \times \begin{bmatrix} 3.83 & 0 \\ 0 & 3.38 \end{bmatrix} \times \begin{bmatrix} 0.34 & -0.37 & -0.58 & 0.64 \\ -0.81 & 0.18 & 0.03 & 0.56 \end{bmatrix}$$

$$= \begin{bmatrix} -1.2 & 0.91 & 1.28 & -1.1 \\ 0.36 & -0.65 & -1.11 & 1.47 \\ 0.02 & 0.56 & 1.07 & -1.71 \\ 2.32 & -0.89 & -0.83 & -0.46 \\ -1.49 & 0.08 & -0.41 & 1.79 \end{bmatrix}$$

Part f

Note that since here we are dealing with a sample covariance matrix, we have a symmetric non-negative definite matrix, so an eigen analysis will also yield all the information we need for a singular value decomposition, and thus we can easily find the best 1 and 2 component approximations.

We begin with the matrix:

$$\frac{1}{5}\tilde{X}'\tilde{X} = \begin{bmatrix} 2 & -0.6 & -0.4 & -0.2 \\ -0.6 & 0.56 & 0.76 & -0.36 \\ -0.4 & 0.76 & 1.36 & -0.76 \\ -0.2 & -0.36 & -0.76 & 2.16 \end{bmatrix}$$

An eigen decomposition yields:

$$\text{Eigenvalues} = D = \begin{bmatrix} 2.94 \\ 2.29 \\ 0.81 \\ 0.05 \end{bmatrix}, \quad \text{Eigenvectors} = U = V = \begin{bmatrix} 0.34 & 0.81 & -0.45 & -0.18 \\ -0.37 & -0.18 & -0.26 & -0.88 \\ -0.58 & -0.03 & -0.68 & 0.45 \\ 0.64 & -0.56 & -0.52 & 0 \end{bmatrix}$$

Thus, a best 1 component approximation will be of the form:

$$\frac{1}{5}\tilde{X}'\tilde{X}^{*1} = \tilde{U}_1 diag(\tilde{d}_1)\tilde{V}_1' = \begin{bmatrix} 0.34 \\ -0.37 \\ -0.58 \\ 0.64 \end{bmatrix} \times \begin{bmatrix} 2.94 \end{bmatrix} \times \begin{bmatrix} 0.34 & -0.37 & -0.58 & 0.64 \end{bmatrix}$$

$$= \begin{bmatrix} 0.35 & -0.37 & -0.58 & 0.65 \\ -0.37 & 0.4 & 0.62 & -0.7 \\ -0.58 & 0.62 & 0.97 & -1.09 \\ 0.65 & -0.7 & -1.09 & 1.22 \end{bmatrix}$$

A best 2 component approximation will be of the form:

$$\frac{1}{5}\tilde{X}'\tilde{X}^{*2} = \tilde{U}_2 diag(\tilde{d}_1, \tilde{d}_2)\tilde{V}_2' = \begin{bmatrix} 0.34 & 0.81 \\ -0.37 & -0.18 \\ -0.58 & -0.03 \\ 0.64 & -0.56 \end{bmatrix} \times \begin{bmatrix} 2.94 & 0 \\ 0 & 2.29 \end{bmatrix} \times \begin{bmatrix} 0.34 & -0.37 & -0.58 & 0.64 \\ 0.81 & -0.18 & -0.03 & -0.56 \end{bmatrix}$$

$$= \begin{bmatrix} 1.84 & -0.7 & -0.64 & -0.39 \\ -0.7 & 0.47 & 0.64 & -0.47 \\ -0.64 & 0.64 & 0.97 & -1.05 \\ -0.39 & -0.47 & -1.05 & 1.94 \end{bmatrix}$$

Part d - standardized

Note that

$$\tilde{\tilde{X}} = \begin{bmatrix} -0.63 & 0.96 & 1.38 & -0.49 \\ 0.63 & -0.24 & -0.15 & 1.34 \\ 0 & 0.96 & 0.61 & -1.1 \\ 1.26 & -1.43 & -0.92 & -0.49 \\ -1.26 & -0.24 & -0.92 & 0.73 \end{bmatrix}$$

The Singular Value Decomposition of $\tilde{\tilde{X}}$ is:

$$\tilde{\tilde{U}} = \begin{bmatrix} -0.6 & 0 & 0.14 & -0.65 \\ 0.3 & 0.22 & 0.79 & 0.2 \\ -0.44 & -0.34 & -0.16 & 0.68 \\ 0.58 & -0.58 & -0.24 & -0.27 \\ 0.16 & 0.7 & -0.53 & 0.04 \end{bmatrix}, \quad \tilde{\tilde{D}} = \begin{bmatrix} 3.04 & 0 & 0 & 0 \\ 0 & 2.14 & 0 & 0 \\ 0 & 0 & 1.39 & 0 \\ 0 & 0 & 0 & 0.49 \end{bmatrix}$$

$$\tilde{\tilde{V}} = \begin{bmatrix} 0.36 & -0.69 & 0.56 & 0.29 \\ -0.63 & 0.13 & 0.19 & 0.74 \\ -0.6 & -0.17 & 0.49 & -0.61 \\ 0.33 & 0.69 & 0.64 & 0 \end{bmatrix}$$

Therefore, the principal component directions will be the columns of $\tilde{\tilde{V}}$, the principal components will be the inner products $z_j = \langle x_i, v_j \rangle$, the columns of:

$$z = \begin{bmatrix} -1.82 & -0.01 & 0.19 & -0.32 \\ 0.92 & 0.48 & 1.09 & 0.1 \\ -1.34 & -0.73 & -0.22 & 0.33 \\ 1.75 & -1.24 & -0.33 & -0.13 \\ 0.49 & 1.5 & -0.73 & 0.02 \end{bmatrix}$$

The "loadings" of the first principal component are the first column of the matrix $\tilde{\tilde{V}}$,

$$\tilde{V}_1 = \begin{bmatrix} 0.36 \\ -0.63 \\ -0.6 \\ 0.33 \end{bmatrix}$$

Part e - standardized

The best $rank = 1$ approximation to $\tilde{\tilde{X}}$ will be $\tilde{\tilde{X}}^{*1} = \tilde{\tilde{U}}_1 diag(\tilde{\tilde{d}}_1), \tilde{\tilde{V}}_1$, where $\tilde{\tilde{U}}_1$ is a matrix with first column of $\tilde{\tilde{U}}$ and $\tilde{\tilde{V}}_1$ is a matrix with the first column of $\tilde{\tilde{V}}$:

$$\tilde{\tilde{X}}^{*1} = \tilde{\tilde{U}}_1 diag(\tilde{\tilde{d}}_1)\tilde{\tilde{V}}_1' = \begin{bmatrix} -0.6 \\ 0.3 \\ -0.44 \\ 0.58 \\ 0.16 \end{bmatrix} \times \begin{bmatrix} 3.04 \end{bmatrix} \times \begin{bmatrix} 0.36 & -0.63 & -0.6 & 0.33 \end{bmatrix}$$

$$= \begin{bmatrix} -0.65 & 1.15 & 1.09 & -0.61 \\ 0.33 & -0.58 & -0.55 & 0.31 \\ -0.48 & 0.85 & 0.8 & -0.45 \\ 0.63 & -1.11 & -1.05 & 0.58 \\ 0.18 & -0.31 & -0.29 & 0.16 \end{bmatrix}$$

The best $rank = 2$ approximation to $\tilde{\tilde{X}}$ will be $\tilde{\tilde{X}}^{*2} = \tilde{\tilde{U}}_2 diag(\tilde{\tilde{d}}_1, \tilde{\tilde{d}}_2)\tilde{\tilde{V}}_2'$, where $\tilde{\tilde{U}}_2$ is a matrix with the first two columns of $\tilde{\tilde{U}}$, and $\tilde{\tilde{V}}_2$ is a matrix with the first two columns of $\tilde{\tilde{V}}$:

$$\tilde{\tilde{X}}^{*2} = \tilde{\tilde{U}}_2 diag(\tilde{\tilde{d}}_1, \tilde{\tilde{d}}_2)\tilde{\tilde{V}}_2' = \begin{bmatrix} -0.6 & 0 \\ 0.3 & 0.22 \\ -0.44 & -0.34 \\ 0.58 & -0.58 \\ 0.16 & 0.7 \end{bmatrix} \times \begin{bmatrix} 3.04 & 0 \\ 0 & 2.14 \end{bmatrix} \times \begin{bmatrix} 0.36 & -0.63 & -0.6 & 0.33 \\ -0.69 & 0.13 & -0.17 & 0.69 \end{bmatrix}$$

$$= \begin{bmatrix} -0.65 & 1.15 & 1.09 & -0.61 \\ 0 & -0.52 & -0.63 & 0.64 \\ 0.03 & 0.75 & 0.93 & -0.95 \\ 1.49 & -1.27 & -0.84 & -0.27 \\ -0.86 & -0.12 & -0.55 & 1.2 \end{bmatrix}$$

Part f - standardized

Note that since here we are dealing with a sample covariance matrix, we have a symmetric non-negative definite matrix, so an eigen analysis will also yield all the information we need for a singular value decomposition, and thus we can easily find the best 1 and 2 component approximations.

We begin with the matrix:

$$\frac{1}{5}\tilde{\tilde{X}}'\tilde{\tilde{X}} = \begin{bmatrix} 0.8 & -0.45 & -0.19 & -0.08 \\ -0.45 & 0.8 & 0.7 & -0.26 \\ -0.19 & 0.7 & 0.8 & -0.35 \\ -0.08 & -0.26 & -0.35 & 0.8 \end{bmatrix}$$

An eigen decomposition yields:

$$\text{Eigenvalues} = D = \begin{bmatrix} 1.85 \\ 0.91 \\ 0.39 \\ 0.05 \end{bmatrix}, \quad \text{Eigenvectors} = U = V = \begin{bmatrix} 0.36 & -0.69 & -0.56 & 0.29 \\ -0.63 & 0.13 & -0.19 & 0.74 \\ -0.6 & -0.17 & -0.49 & -0.61 \\ 0.33 & 0.69 & -0.64 & 0 \end{bmatrix}$$

Thus, a best 1 component approximation will be of the form:

$$\frac{1}{5}\tilde{\tilde{X}}'\tilde{\tilde{X}}^{*1} = \tilde{\tilde{U}}_1 diag(\tilde{\tilde{d}}_1)\tilde{\tilde{V}}_1' = \begin{bmatrix} 0.36 \\ -0.63 \\ -0.6 \\ 0.33 \end{bmatrix} \times \begin{bmatrix} 1.85 \end{bmatrix} \times \begin{bmatrix} 0.36 & -0.63 & -0.6 & 0.33 \end{bmatrix}$$

$$= \begin{bmatrix} 0.24 & -0.42 & -0.4 & 0.22 \\ -0.42 & 0.74 & 0.7 & -0.39 \\ -0.4 & 0.7 & 0.66 & -0.37 \\ 0.22 & -0.39 & -0.37 & 0.21 \end{bmatrix}$$

A best 2 component approximation will be of the form:

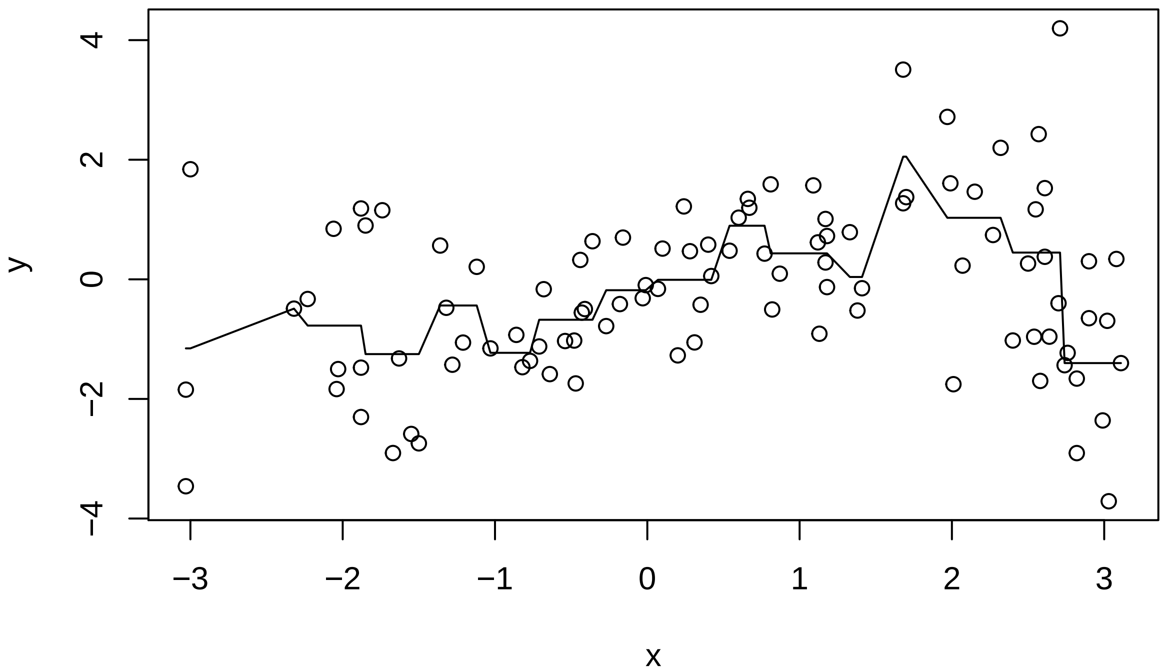
$$\frac{1}{5}\tilde{\tilde{X}}'\tilde{\tilde{X}}^{*2} = \tilde{\tilde{U}}_2 diag(\tilde{\tilde{d}}_1, \tilde{\tilde{d}}_2)\tilde{\tilde{V}}_2' = \begin{bmatrix} 0.36 & -0.69 \\ -0.63 & 0.13 \\ -0.6 & -0.17 \\ 0.33 & 0.69 \end{bmatrix} \times \begin{bmatrix} 1.85 & 0 \\ 0 & 0.91 \end{bmatrix} \times \begin{bmatrix} 0.36 & -0.63 & -0.6 & 0.33 \\ -0.69 & 0.13 & -0.17 & 0.69 \end{bmatrix}$$

$$= \begin{bmatrix} 0.68 & -0.5 & -0.29 & -0.21 \\ -0.5 & 0.76 & 0.68 & -0.31 \\ -0.29 & 0.68 & 0.69 & -0.48 \\ -0.21 & -0.31 & -0.48 & 0.64 \end{bmatrix}$$

Problem 2

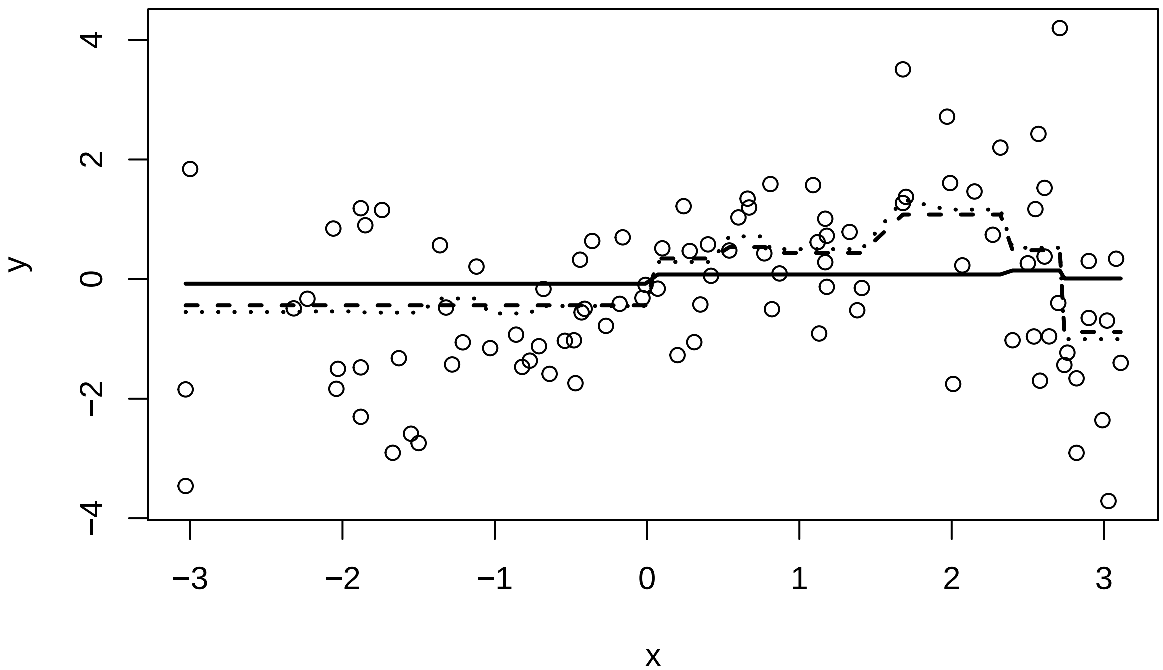
Part a

A scatterplot of the data and the fitted values from the OLS model are shown below.



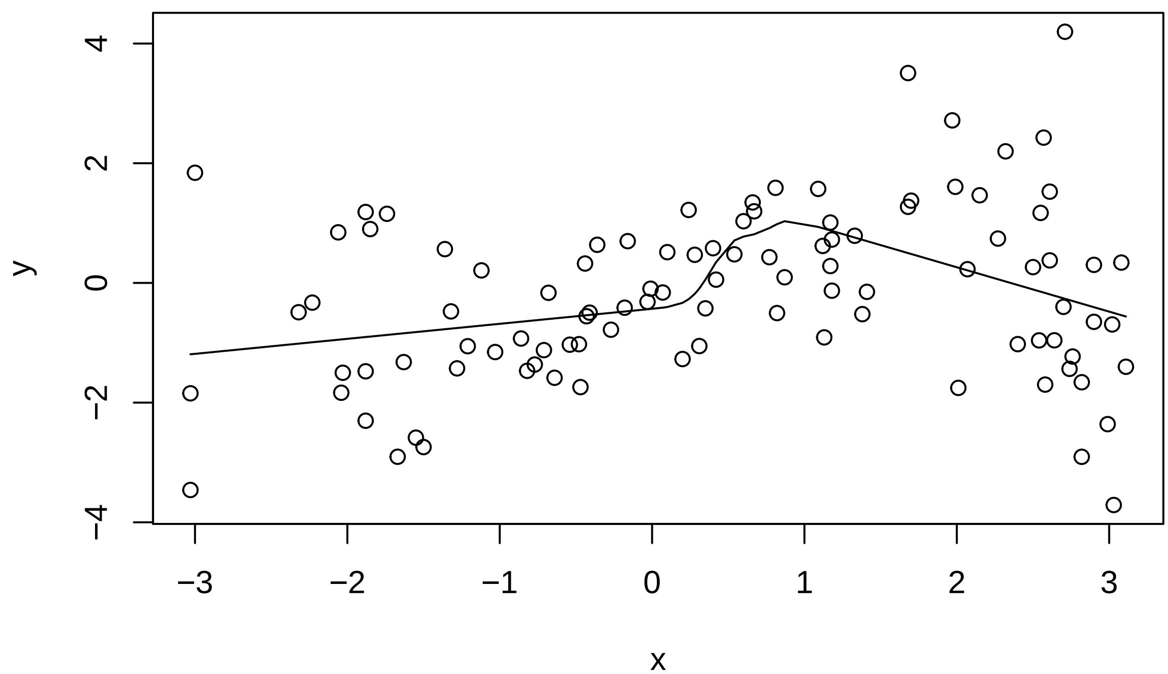
Part b

The solid line in the plot below gives (approximately) the fitted values when there are 2 nonzero coefficients, the dashed line are fitted values when there are 4 nonzero coefficients, and the dotted line are fitted values for 8 nonzero coefficients.



y
x
4
2
0
−2
−4
−3
−2
−1
0
1
2
3

Problem 3



Problem 4

Part a

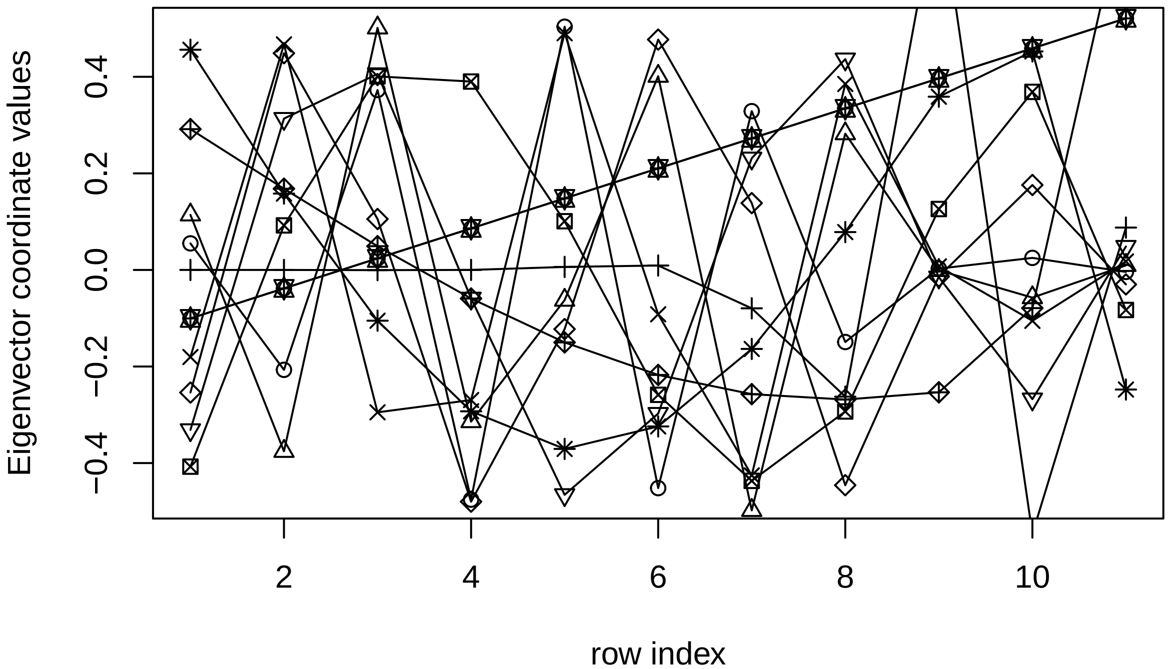
```
## [1] "Matrix H"

##       [,1] [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
##  [1,]    1  0.0 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##  [2,]    1  0.1 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##  [3,]    1  0.2 0.008 0.001 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##  [4,]    1  0.3 0.027 0.008 0.001 0.000 0.000 0.000 0.000 0.000 0.000
##  [5,]    1  0.4 0.064 0.027 0.008 0.001 0.000 0.000 0.000 0.000 0.000
##  [6,]    1  0.5 0.125 0.064 0.027 0.008 0.001 0.000 0.000 0.000 0.000
##  [7,]    1  0.6 0.216 0.125 0.064 0.027 0.008 0.001 0.000 0.000 0.000
##  [8,]    1  0.7 0.343 0.216 0.125 0.064 0.027 0.008 0.001 0.000 0.000
##  [9,]    1  0.8 0.512 0.343 0.216 0.125 0.064 0.027 0.008 0.001 0.000
## [10,]    1  0.9 0.729 0.512 0.343 0.216 0.125 0.064 0.027 0.008 0.001
## [11,]    1  1.0 1.010 0.738 0.520 0.350 0.222 0.130 0.068 0.030 0.010

## [1] "Matrix Omega"

##       [,1] [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
##  [1,]    0    0 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##  [2,]    0    0 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000
##  [3,]    0    0 4.704 3.744 2.820 1.968 1.224 0.624 0.204 0.000 0.048
##  [4,]    0    0 3.744 3.024 2.310 1.632 1.026 0.528 0.174 0.000 0.042
```

```
## [5,]    0    0 2.820 2.310 1.800 1.296 0.828 0.432 0.144 0.000 0.036
## [6,]    0    0 1.968 1.632 1.296 0.960 0.630 0.336 0.114 0.000 0.030
## [7,]    0    0 1.224 1.026 0.828 0.630 0.432 0.240 0.084 0.000 0.024
## [8,]    0    0 0.624 0.528 0.432 0.336 0.240 0.144 0.054 0.000 0.018
## [9,]    0    0 0.204 0.174 0.144 0.114 0.084 0.054 0.024 0.000 0.012
## [10,]   0    0 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.000 0.006
## [11,]   0    0 0.048 0.042 0.036 0.030 0.024 0.018 0.012 0.006 0.000
```

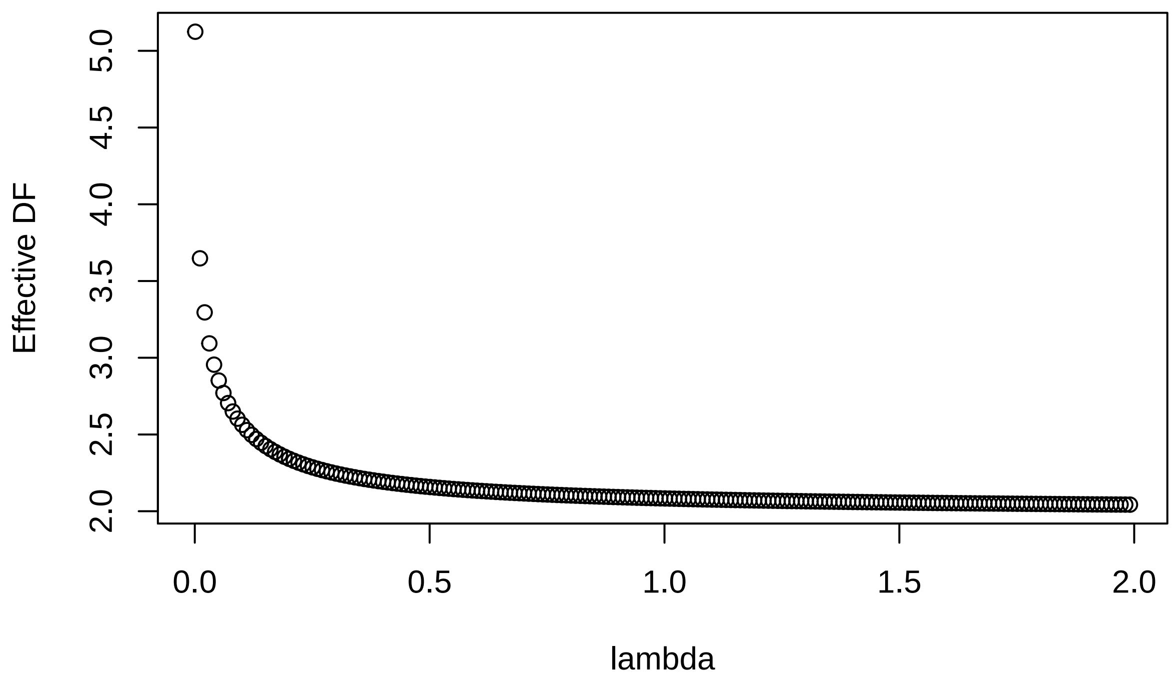
```
## [1] "Matrix K"
```

```
##              [,1]          [,2]          [,3]          [,4]         [,5]
##  [1,] 1607.6951262 -3646.1707571  2.584683e+03   -692.561356    185.56240
##  [2,] -3646.1707571  9877.0245425 -9.508098e+03   4155.368138  -1113.37438
##  [3,] 2584.6830284 -9508.0981702  1.403239e+04 -10621.472553   4453.49753
##  [4,]  -692.5613564  4155.3681382 -1.062147e+04  14330.522073 -10700.61574
##  [5,]   185.5623971 -1113.3743825  4.453498e+03 -10700.615737  14348.96542
##  [6,]   -49.6882320   298.1293918 -1.192518e+03   4471.940877 -10695.24594
##  [7,]    13.1905308   -79.1431848  3.165727e+02  -1187.147772   4432.01835
##  [8,]    -3.0738912    18.4433474 -7.377339e+01    276.650211  -1032.82746
##  [9,]     0.1774497    -1.0646979  4.258792e+00    -15.970469     59.62308
## [10,]     0.2222482    -1.3334891  5.333956e+00    -20.002336     74.67539
## [11,]    -0.0365436     0.2192616 -8.770464e-01      3.288924    -12.27865
##             [,6]         [,7]         [,8]          [,9]         [,10]
##  [1,]    -49.68823     13.19053    -3.073891     0.1774497     0.2222482
##  [2,]    298.12939    -79.14318    18.443347    -1.0646979    -1.3334891
##  [3,]  -1192.51757    316.57274   -73.773390     4.2587917     5.3339563
##  [4,]   4471.94088  -1187.14777   276.650211   -15.9704689   -20.0023361
##  [5,] -10695.24594   4432.01835 -1032.827456    59.6230838    74.6753880
##  [6,]  14309.04289 -10540.92563  3854.659612  -222.5218665  -278.6992159
##  [7,] -10540.92563  13731.68415 -8385.810993   830.4643821  1040.1214758
##  [8,]   3854.65961  -8385.81099  5688.584360  2900.6643383 -3881.7866872
##  [9,]   -222.52187    830.46438  2900.664338 -8602.1481175  6020.8472547
## [10,]   -278.69922   1040.12148 -3881.786687  6020.8472547 -3504.2529772
## [11,]     45.82568   -171.02406   638.270548  -974.3301494   544.8743825
##             [,11]
##  [1,]   -0.0365436
##  [2,]    0.2192616
##  [3,]   -0.8770464
##  [4,]    3.2889242
##  [5,]  -12.2786502
##  [6,]   45.8256766
##  [7,] -171.0240560
##  [8,]  638.2705476
##  [9,] -974.3301494
## [10,]  544.8743825
## [11,]  -73.9323467
```

Part b


Eigenvector coordinate values
−0.4
−0.2
0.0
0.2
0.4
2
4
6
8
10
row index

Part c




```
## [1] 5.000493
## [1] 4.001968
## [1] 2.998563
## [1] 2.500648
```

The sequence of effective degrees of freedom $2.5, 3, 4, 5$ approximately corresponds to the sequence of λ values $0.0012, 0.0059, 0.0375, 0.12$.

Problem 5

Part a

In order to compute and plot effective degrees of freedom, we need to compute the matrix

$$L_\lambda = \begin{bmatrix} l'(x_1) \\ \vdots \\ l'(x_{11}) \end{bmatrix}$$

where $l'(x_i) = (1, x_i)(\mathbf{B}'\mathbf{W}(x_i)\mathbf{B})^{-1}\mathbf{B}'\mathbf{W}(x_i)$, where

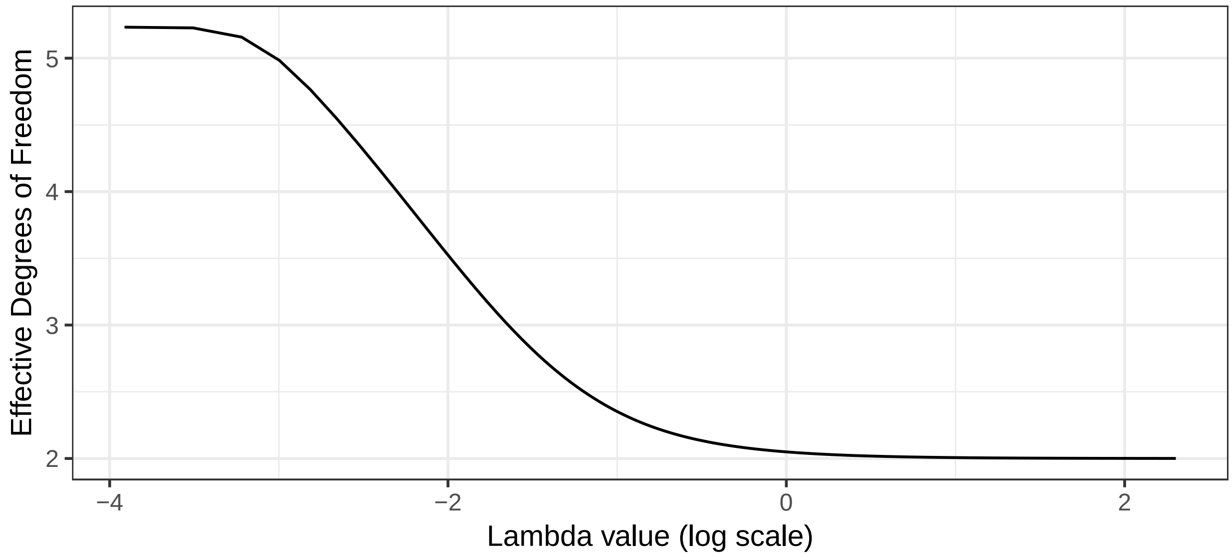


Figure 1: A plot of effective degrees of freedom as a function of Lambda values for Problem 5.

$$\mathbf{B} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_{11} \end{bmatrix}, \quad \mathbf{W} = diag(K_\lambda(x_i, x_1), \ldots, K_\lambda(x_i, x_N))$$

A plot can be seen in Figure 1. The corresponding table for specific degrees of freedom can be seen in Table 2.

Table 2: Values of λ for effective degrees of freedom.

Effective DF	λ value
2.5	0.2997197
3	0.1898298
4	0.0999199
5	0.04997

Part b

First, note that we are working with $\lambda = 0.0999$, and our L_λ matrix is:

```
##        [,1]  [,2]  [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11]
##  [1,]  0.97  0.06 -0.03 0.00 0.00 0.00 0.00 0.00 0.00  0.00  0.00
##  [2,]  0.31  0.43  0.22 0.04 0.00 0.00 0.00 0.00 0.00  0.00  0.00
##  [3,] -0.03  0.35  0.43 0.21 0.04 0.00 0.00 0.00 0.00  0.00  0.00
##  [4,] -0.15  0.16  0.40 0.39 0.17 0.03 0.00 0.00 0.00  0.00  0.00
##  [5,] -0.19  0.03  0.25 0.40 0.35 0.15 0.03 0.00 0.00  0.00  0.00
##  [6,] -0.20 -0.04  0.13 0.28 0.38 0.31 0.12 0.02 0.00  0.00  0.00
##  [7,] -0.19 -0.07  0.05 0.17 0.29 0.35 0.27 0.11 0.02  0.00  0.00
##  [8,] -0.18 -0.09  0.01 0.10 0.20 0.29 0.33 0.24 0.09  0.02  0.00
##  [9,] -0.17 -0.10 -0.02 0.06 0.13 0.21 0.28 0.30 0.22  0.08  0.01
## [10,] -0.17 -0.10 -0.04 0.02 0.09 0.15 0.21 0.27 0.29  0.20  0.07
## [11,] -0.17 -0.11 -0.06 0.00 0.06 0.11 0.17 0.23 0.28  0.29  0.20
```

And our S_λ matrix is:

```
##        [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
##  [1,]  0.60  0.35  0.15  0.04 -0.02 -0.04 -0.04 -0.03 -0.01  0.00  0.00
##  [2,]  0.35  0.30  0.21  0.12  0.05  0.01 -0.01 -0.01 -0.02 -0.01  0.00
##  [3,]  0.15  0.21  0.24  0.20  0.13  0.07  0.03  0.00 -0.02 -0.02  0.00
##  [4,]  0.04  0.12  0.20  0.24  0.21  0.14  0.08  0.03 -0.01 -0.03  0.00
##  [5,] -0.02  0.05  0.13  0.21  0.24  0.21  0.14  0.07  0.01 -0.03  0.00
##  [6,] -0.04  0.01  0.07  0.14  0.21  0.24  0.20  0.13  0.05 -0.01  0.00
##  [7,] -0.04 -0.01  0.03  0.08  0.14  0.20  0.24  0.20  0.12  0.05 -0.01
##  [8,] -0.03 -0.01  0.00  0.03  0.07  0.13  0.20  0.24  0.22  0.17 -0.02
##  [9,] -0.01 -0.02 -0.02 -0.01  0.01  0.05  0.12  0.22  0.32  0.35 -0.02
## [10,]  0.00 -0.01 -0.02 -0.03 -0.03 -0.01  0.05  0.17  0.35  0.42  0.13
## [11,]  0.00  0.00  0.00  0.00  0.00  0.00 -0.01 -0.02 -0.02  0.13  0.92
```

Thus, the 11x11 matrix difference is:

```
##        [,1]  [,2]  [,3]  [,4]  [,5]  [,6]  [,7]  [,8]  [,9] [,10] [,11]
##  [1,] -0.37  0.28  0.18  0.04 -0.02 -0.04 -0.04 -0.03 -0.01  0.00  0.00
##  [2,]  0.04 -0.13  0.00  0.08  0.05  0.01 -0.01 -0.01 -0.02 -0.01  0.00
##  [3,]  0.18 -0.14 -0.19 -0.01  0.09  0.07  0.03  0.00 -0.02 -0.02  0.00
##  [4,]  0.19 -0.03 -0.20 -0.15  0.03  0.11  0.07  0.03 -0.01 -0.03  0.00
##  [5,]  0.17  0.03 -0.11 -0.19 -0.11  0.06  0.11  0.07  0.01 -0.03  0.00
##  [6,]  0.16  0.05 -0.05 -0.14 -0.17 -0.07  0.08  0.11  0.05 -0.01  0.00
##  [7,]  0.16  0.06 -0.02 -0.10 -0.15 -0.15 -0.03  0.09  0.11  0.05 -0.01
##  [8,]  0.16  0.07 -0.01 -0.08 -0.13 -0.16 -0.13  0.00  0.13  0.15 -0.02
##  [9,]  0.16  0.08  0.00 -0.07 -0.12 -0.16 -0.15 -0.08  0.10  0.27 -0.03
## [10,]  0.17  0.09  0.01 -0.06 -0.12 -0.16 -0.16 -0.10  0.06  0.22  0.06
## [11,]  0.17  0.11  0.06  0.00 -0.06 -0.12 -0.18 -0.24 -0.30 -0.16  0.72
```

The described plot can be seen in Figure 2.

Problem 6

Part a

The plots for this problem can be seen in Figure 3.

Problem 7

A scatterplot matrix for the five different methods and the response can be seen in Figure 4.

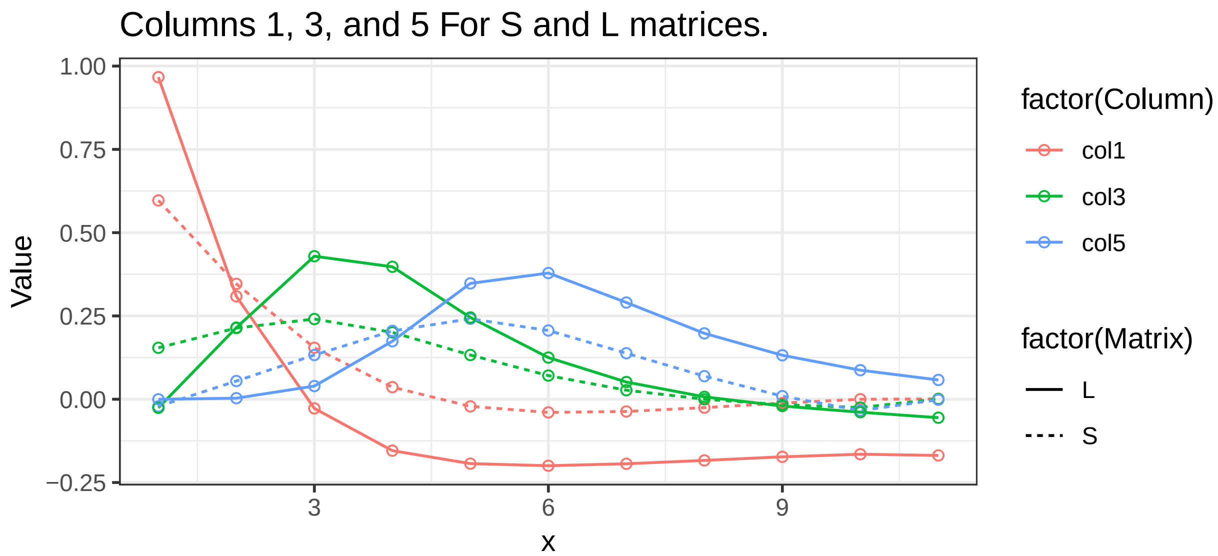


Figure 2: Plot comparing the first, third, and fifth columns of the S_lambda and L_lambda matrices for 4 effective degrees of freedom. This if for problem 5 part b.

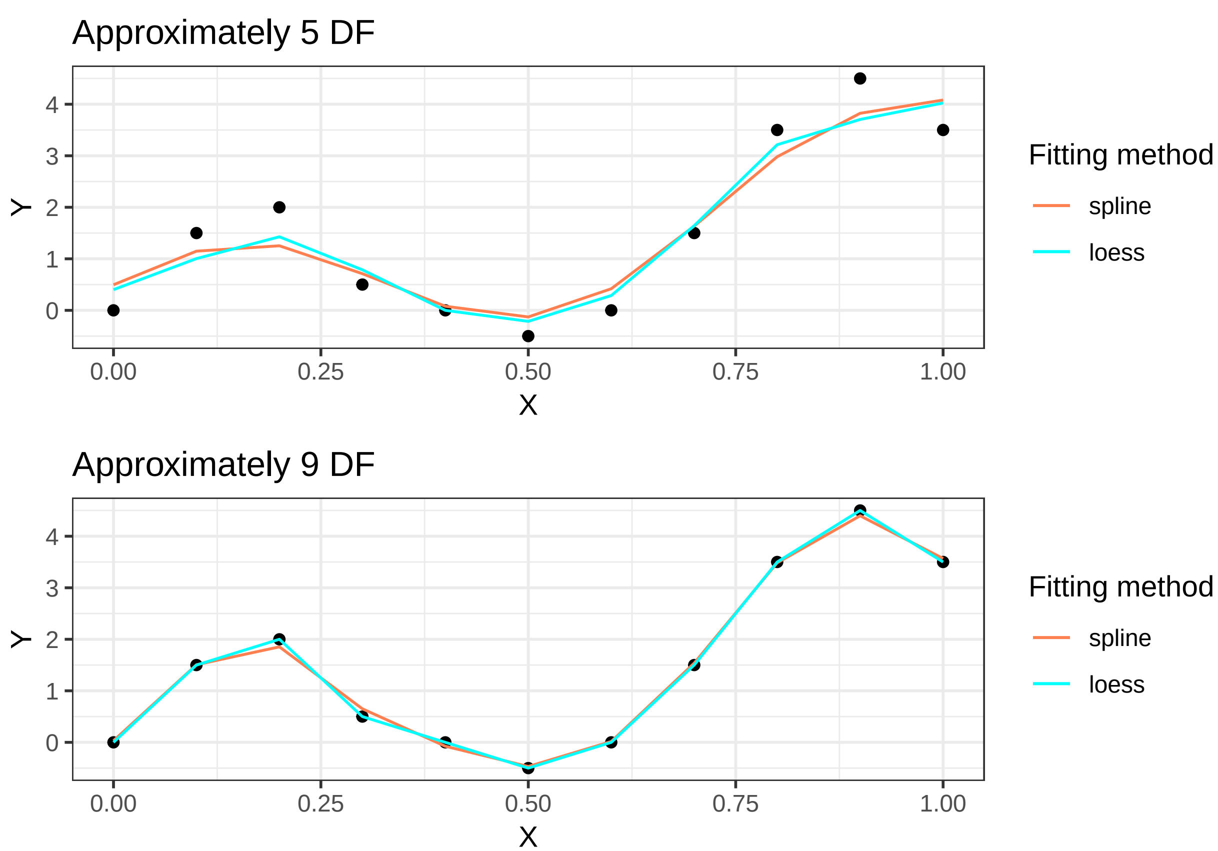


Figure 3: Plots for effective degrees of freedom 5 (top) and 9 (bottom) for Problem 6(a).

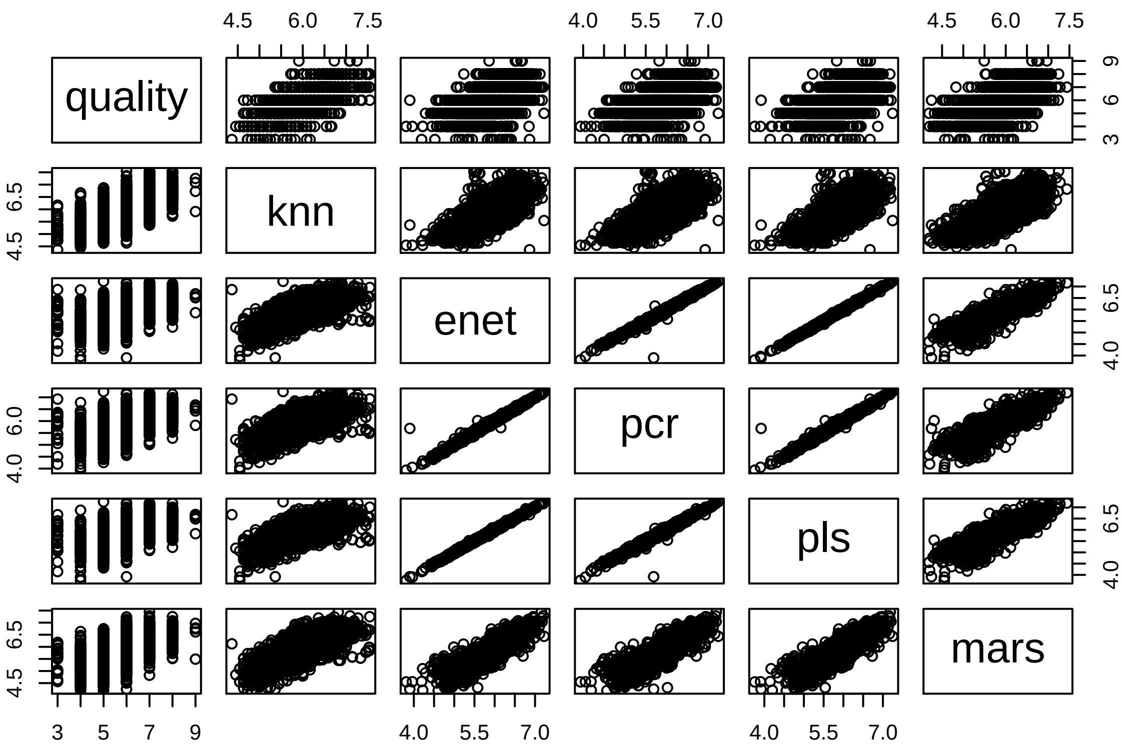


Figure 4: A scatterplot matrix for the five different prediction methods in Problem 7.

A correlation matrix can be seen below:

```
##           quality       knn      enet       pcr       pls      mars
## quality 1.0000000 0.6940572 0.5303217 0.5251641 0.5294362 0.5751068
## knn     0.6940572 1.0000000 0.7090516 0.7054792 0.7069169 0.7519074
## enet    0.5303217 0.7090516 1.0000000 0.9948869 0.9982166 0.8923727
## pcr     0.5251641 0.7054792 0.9948869 1.0000000 0.9932919 0.8778578
## pls     0.5294362 0.7069169 0.9982166 0.9932919 1.0000000 0.8900391
## mars    0.5751068 0.7519074 0.8923727 0.8778578 0.8900391 1.0000000
```

Problem 8

Part a

Consider the Gram-Schmidt process using the L_2 inner product $\int f(x)g(x)dx$ on functions $f_c(x) = e^{-\frac{(c-x)^2}{2}}$ for $c = x_1, x_2, x_3$. Define

$$
\begin{aligned}
u_1(x) &= f_{x_1}(x) / \left[\int f_{x_1}^2(x)dx \right]^{1/2} && (1) \\
&= e^{-\frac{(x_1-x)^2}{2}} / \left[\int e^{-(x_1-x)^2} dx \right]^{1/2} && (2) \\
&= \pi^{1/4} e^{-\frac{(x_1-x)^2}{2}}. && (3)
\end{aligned}
$$

Next, define

$$
\begin{aligned}
u_2(x) &= \frac{f_{x_2}(x) - \int f_{x_2}(x)u_1(x)dxu_1(x)}{\left[\int (f_{x_2}(x) - \int f_{x_2}(x)u_1(x)dxu_1(x))^2\right]^{1/2}} && (4) \\
&= \frac{f_{x_2}(x) - (\pi^{1/4} e^{\frac{1}{4}(x_1+x_2)^2 - 2x_1x_2}) \int e^{-\frac{1}{2*1/2}(x-1/2(x_1+x_2)^2)} dx * u_1(x)}{\left[\int (f_{x_2}(x) - (\pi^{1/4} e^{\frac{1}{4}(x_1+x_2)^2 - 2x_1x_2}) \int e^{-\frac{1}{2*1/2}(x-1/2(x_1+x_2)^2)} dx * u_1(x))^2\right]^{1/2}} && (5) \\
&= \frac{f_{x_2}(x) - (\pi^{1/4} e^{\frac{1}{4}(x_1+x_2)^2 - 2x_1x_2}) \pi^{1/2} \pi^{1/4} e^{-\frac{(x_1-x)^2}{2}}}{\left[\int (f_{x_2}(x) - (\pi^{1/4} e^{\frac{1}{4}(x_1+x_2)^2 - 2x_1x_2}) \pi^{1/2} \pi^{1/4} e^{-\frac{(x_1-x)^2}{2}})^2\right]^{1/2}} && (6) \\
&= \frac{e^{-(x_2-x)^2}/2}{\pi^{1/2} - 2q(x_1,x_2)\sqrt{2\pi} + q(x_1,x_2)}, && (7)
\end{aligned}
$$

where $q(c_1, c_2) = \pi e^{\frac{1}{4}(c_1+c_2)^2 - 2c_1c_2}$.

Following the same procedure for the third function $u_3(x)$, we get

$$
\begin{aligned}
u_3(x) &= \frac{f_{x_3}(x) - \int f_{x_3}(x)u_2(x)dx * u_2(x) - \int f_{x_3}(x)u_1(x)dx * u_1(x)}{\left[\int (f_{x_3}(x) - \int f_{x_3}(x)u_2(x)dx * u_2(x) - \int f_{x_3}(x)u_1(x)dx * u_1(x))^2\right]^{1/2}} && (8) \\
&= \frac{f_{x_3}(x) - \int f_{x_3}(x)u_2(x)dx * u_2(x) - \int f_{x_3}(x)u_1(x)dx * u_1(x)}{v(x_1,x_2,x_3)} && (9) \\
&= \frac{f_{x_3}(x) - \frac{\pi^{-1/2} q(x_2,x_3) e^{-(x_2-x)^2/2}}{\pi^{1/2} + 2q(x_2,x_3)\sqrt{2\pi} + q(x_2,x_3)} - 1(x_1,x_3)e^{-(x_1-x)^2/2}}{v(x_1,x_2,x_3)}, && (10)
\end{aligned}
$$

where $v(x_1,x_2,x_3) = \pi^{1/2} + \left(\frac{q(x_2,x_3)}{\pi^{1/2}+2q(x_2,x_3)\sqrt{2\pi}+q(x_2,x_3)}\right)^2 + q(x_1,x_3)^2\pi^{1/2} - 2\frac{\pi^{-1/2}q(x_2,x_3)q(x_1,x_3)}{\pi^{1/2}+2q(x_2,x_3)\sqrt{2\pi}+q(x_2,x_3)}\pi^{-1/2}q(x_1,x_2) - 2\frac{\pi^{-1/2}q(x_2,x_3)}{\pi^{1/2}+2q(x_2,x_3)\sqrt{2\pi}+q(x_2,x_3)}\pi^{-1/2}q(x_2,x_3) - 2q(x_1,x_3)\pi^{-1/2}q(x_1,x_3)$.

If we consider the kernel inner product, the inner product expressions are much nicer. Using the kernel inner product, we wind up with

$$u_1(x) = f_{x_1}(x) \tag{11}$$

$$u_2(x) = \frac{f_{x_2}(x) - K(x_1,x_2)u_1(x)}{|1 - K(x_1,x_2)|} \tag{12}$$

$$u_3(x) = \frac{f_{x_3}(x) - \langle f_{x_3}(x), u_2(x)\rangle u_2(x) - \langle f_{x_3}(x), u_1(x)\rangle u_1(x)}{\langle f_{x_3}(x) - \langle f_{x_3}(x), u_2(x)\rangle u_2(x) - \langle f_{x_3}(x), u_1(x)\rangle u_1(x), f_{x_3}(x) - \langle f_{x_3}(x), u_2(x)\rangle u_2(x) - \langle f_{x_3}(x), u_1(x)\rangle u_1(x)\rangle}. \tag{13}$$

These are not the same functions, as the inner products are different. However, the inner products essentially just control constants that amount to centering and scaling factors.

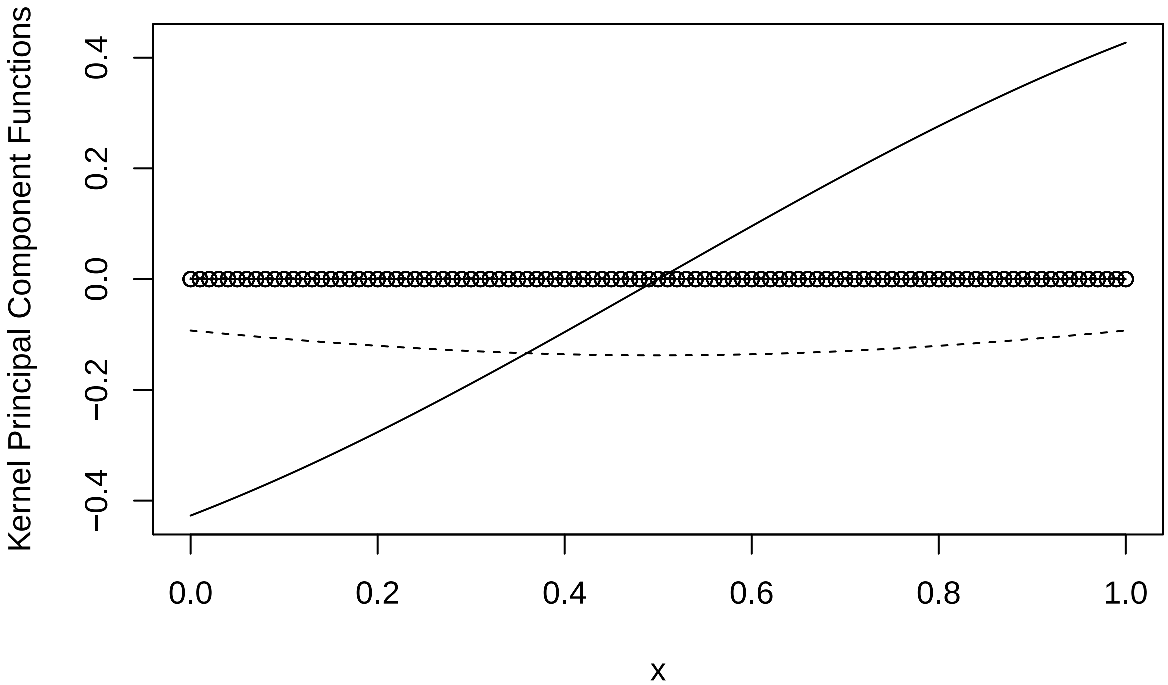
Part b

```
## [1] "Centered Gram Matrix"
```

```
##               [,1]         [,2]         [,3]         [,4]         [,5]
##  [1,]  0.209980241  0.169676192  0.126387671  0.081291166  0.035645321
##  [2,]  0.169676192  0.139347183  0.105884948  0.070175829  0.033209927
##  [3,]  0.126387671  0.105884948  0.082397754  0.056514920  0.028936404
##  [4,]  0.081291166  0.070175829  0.056514920  0.040607127  0.022854897
##  [5,]  0.035645321  0.033209927  0.028936404  0.022854897  0.015077708
##  [6,] -0.009267458 -0.003964544  0.000441877  0.003747755  0.005796851
##  [7,] -0.052200814 -0.040290652 -0.028145923 -0.016160101 -0.004723619
##  [8,] -0.092001777 -0.074752633 -0.056000657 -0.036276526 -0.016160101
##  [9,] -0.127661965 -0.106422993 -0.082332035 -0.056000657 -0.028145923
## [10,] -0.158359477 -0.134503780 -0.106422993 -0.074752633 -0.040290652
## [11,] -0.183489099 -0.158359477 -0.127661965 -0.092001777 -0.052200814
##               [,6]         [,7]         [,8]         [,9]        [,10]
##  [1,] -0.009267458 -0.052200814 -0.092001777 -0.127661965 -0.158359477
##  [2,] -0.003964544 -0.040290652 -0.074752633 -0.106422993 -0.134503780
##  [3,]  0.000441877 -0.028145923 -0.056000657 -0.082332035 -0.106422993
##  [4,]  0.003747755 -0.016160101 -0.036276526 -0.056000657 -0.074752633
##  [5,]  0.005796851 -0.004723619 -0.016160101 -0.028145923 -0.040290652
##  [6,]  0.006491037  0.005796851  0.003747755  0.000441877 -0.003964544
##  [7,]  0.005796851  0.015077708  0.022854897  0.028936404  0.033209927
##  [8,]  0.003747755  0.022854897  0.040607127  0.056514920  0.070175829
##  [9,]  0.000441877  0.028936404  0.056514920  0.082397754  0.105884948
## [10,] -0.003964544  0.033209927  0.070175829  0.105884948  0.139347183
## [11,] -0.009267458  0.035645321  0.081291166  0.126387671  0.169676192
##              [,11]
##  [1,] -0.183489099
##  [2,] -0.158359477
##  [3,] -0.127661965
##  [4,] -0.092001777
##  [5,] -0.052200814
##  [6,] -0.009267458
##  [7,]  0.035645321
```

```
##  [8,]  0.081291166
##  [9,]  0.126387671
## [10,]  0.169676192
## [11,]  0.209980241
```

Part c



As the eigenvalues decrease, the kernal principal components get closer to the zero function.

Part d

This was Nate's problem which he was unable to complete.